UDC: 004.8:004.93 https://doi.org/10.15407/jai2024.01.087

O.V. Savchuk[1], O.M. Morgal[2]
[1,2]National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Ukraine
37, Prospect Beresteiskyi, Kyiv, 03056
[1]savchuk_l1@ukr.net
[2]m_olegm@ukr.net
[1]https://orcid.org/0000-0003-3176-7952
[2]https://orcid.org/0000-0003-4654-9499

ON USAGE OF THE NEURAL NETWORK TECHNOLOGIES IN THE IT- STRUCTURE COMPONENTS' DIAGNOSING

Abstract. The idea of using neural network technologes to prove electrophysical diagnostic methods based on the integral physical effects of IT structure components is considered. It is proposed to transform the received information using a discrete Karhunen-Loeve expansion, which gives the minimum root mean square error of packing a priory vectors in multidimensional space. The use of neural networks: MLP, self-organizing (Kohonen Maps) and RBF in MATLAB environment is verified. The best result for microcircuits was obtained using probabilistic RBF-neural networks. A new neural network approach to diagnostics made it possible to perform individual sorting of elements and ststistical evaluation of the IT structure components batch.

Keywords: neural network technologies; diagnostics; components; IT structure.

Introduction

Among the main problems of creating prerequisites for the post-war recovery of the national economy of Ukraine is ensuring its sustainable development based on breakthrough technologies of artificial intelligence and the corresponding increase in the welfare and quality of the population life. In this monograph [1], artificial intelligence (AI) is defined as a function of artificial consciousness, and among the primary issues are: to ensure the safety and reliability of the AI system; create a reliable communication infrastructure using available computing power.

The modern IT infrastructure with its IT structure and components require constant improvement and updating [2,3]. The stochastic nature of this process requires an increase in the reliability of the technical diagnosis, which can be used to determine the necessary data of the studied component.

The decrease in performance can be explained by the presence of hidden or obvious defects, as well as by various destructive processes at the physical level and various malfunctions at the technical level. All these processes lead to the possibility of changing the technical state, which explains the rapid development of technical diagnostic methods by identifying the relevant characteristics of changes in their physical or technical environment [4].

The main tasks in the field of technical diagnostics of IT structure components are:

1) to determine the type of technical condition of the object by the characteristics of changes in their internal technical or physical environment during the performance of working functions;
2) trace or locate a fault or determine the cause of a fault condition;
3) to predict changes in the technical state of the object to determine the probable causes of such changes or to determine the period of time after which processes may begin in it that will lead to undesirable changes for the technical state of the IT structure due to this component.

The use of neural network technologies is a high-tech direction in the theory of reliability and diagnostics of IT structures. The purpose of the article is to choose the type of neural networks, methods of processing accumulated data to increase the reliability of diagnostic results, and their research on examples of the IT structure components diagnostic process implementation, including integrated microcircuits (ICs).

The Problems to Be Solve:

– Use of technical diagnostic methods to obtain a priori diagnostic information about the state of the component of the IT structure;

– Compression of a priori diagnostic information using the Karhunen-Loeve discrete expansion (KLDE);

– Learning and recognizing a posteriori information for the studied components of the representative sample;

– Selecting the type of neural networks to classify and sort the investigated components according to their physical or technical state in the MATLAB software package and its Neural Network Toolbox library.

The direction of solving the problem consists in checking the possibility of using neural network technologies to improve the electrophysical methods of diagnosing components using the example of microcircuits based on integral physical effects in order to increase the reliability of the IT structure.

I.Metodology for Obtaining a Priory Diagnostic Information on IT Structire Compontnts

When using local diagnostic methods, instrumentally measurable information for the most part has the character of an image of the topographic distribution of certain properties of external environments. This allows not only to detect the presence of defects in the component, but also to localize the location of the defects, indicating their size and orientation in space. The disadvantage of these methods is to increase the operational duration and complexity of the diagnostic process [4].

For integrated methods of physical diagnostics and obtaining diagnostic information, integrated physical effects are used, which are observed in technical or physical environments of various nature. These commonly recognized effects include integral nonlinearity effects, inertial effects, and fluctuation effects. The occurrence of these effects is associated with working processes in the investigated component to be diagnosed (CED) and has the same origin of activation.

This fact means that for the integral effects to be observed, the same sources of energy activation are needed, which ensure the efficiency of the CED during its operation. If the input influences are electrical in nature, then for integrated diagnostic methods it is necessary to use electricity of the same level as for working processes. If there are mechanical products, they require mechanical energy, etc. In all these cases, there is no need to apply additional conditions for the protection of service personnel, except for the implementation of safety rules during operation.

The high efficiency of integrated diagnostic methods is their main advantage due to the short period of operational diagnostic tools operation while reducing the operational labor intensity and operational costs of the diagnostic process. Diagnostic digital information signals often have a variety of analog signatures. Thus, these methods are well suited for rapid diagnosis in adverse conditions to detect not only obvious defects that cause malfunctions in CED, but also to detect hidden defects that cause sudden and gradual failures with loss of CED operational status.

Let us consider two features of nonlinearity in relation to physical effects. First, there is no specific measurement (or metric scale) of nonlinearity as a physical quantity that conveys a quantitative idea of the properties of this object. Secondly, in a quantitative sense, nonlinearity is quite ambiguous, since this property cannot be sufficiently identified by a single feature of functional characteristics. As a rule, several nonlinearity functions are used, which will already make up the range of defining variables (polynomials of different degrees, derivatives of different orders, curvature of the first and higher orders, etc.).

For inertia, the corresponding manifestation of inertial effects is related to the transient and impulse characteristics of CED. Their registration is carried out by the activation of physical quantities at the input of the CED in the form of a pulse jump or a double jump. In this case, energy conversion mechanisms in the dissipative subsystems of the physical environment and mechanisms of energy dissipation into the environment begin to operate. Transient integral characteristics reflect local macro-characteristics of the physical environment at time intervals. Thus, manifest and latent defects cause some transient or impulse changes in characteristics.

Consider a model in the form of a black box, which is defined as

$$Y(t) = Y(t) = g(\cdot)U(t), \quad (1.1)$$

where *U(t)* - input function; *Y(t)* - output function; *t* - time.

For physical dynamic objects $g(\cdot)$ is the susceptibility function – the complex physical environment function. The susceptibility function manifests integrated physical effects: non-linearity, inertia (describes dynamics), fluctuations (own noise) [4]. Let us consider the, susceptibility function components:

$$g(\cdot) = const\ [g(\cdot)] + var\ [g(\cdot)]. \quad (1.2)$$

For physical objects the function $g(.)$ is a complex function of environment

$$g(\cdot) = const\ R_0[g(\cdot)] + var\ Re[g(\cdot)] + J\{const\ Im[g(\cdot)] + +var\ Im[g(\cdot)], \quad (1.3)$$

where *Re* - the real part of the function; *Im* -the imaginary part; *cost* - the fixed part; *var* - the variable part.

Integral physical effects of nonlinearity, inertia, and fluctuations are used to obtain diagnostic information during electro physical diagnostics. According to (1.3) and using the Taylor transformation for the obtained characteristics of electrical components, it is possible to describe, say, an integrated circuit by its dynamic resistance as the 1st derivative of a non-linear ampere-volt characteristic.

The general approach to hardware support of technical diagnostic methods is sufficiently described in [4,5]. Synchronous and parallel methodology of intellectual diagnostic devices and sensors which are intellectual aids for accumulation of new diagnostic knowledge.

II.Conversation of Diagnostic Information

The transformation of primary diagnostic information about the state of CED was carried out using the discrete Karhunen-Loeve expansion (DKLE), which is the expansion of the initial ensemble of vectors by the eigenvectors of the covariance matrix [6-8].

Let *X* be an *n*-dimensional random vector [9], then *X* can be exactly represented by a distribution

$$X = \sum_{i=1}^{n} y_i \Phi_i = \Phi Y_i \quad (2.1)$$

where

$$\Phi = [\Phi_1 \dots \Phi_n], \quad (2.2)$$

$$Y = [y_1 \dots y_n]^T. \quad (2.3)$$

The matrix *Φ* is deterministic and consists of *n* linearly independent column vectors, i.e

$$|\Phi| \neq 0. \quad (2.4)$$

Accordingly, linear combinations of the columns of the matrix *Φ* form an *n*-dimensional space that contains *X*. The columns of the matrix *Φ* are called basis vectors. These columns must be orthonormal, i.e

$$\Phi_i^T \Phi_j = \begin{cases} 1, i = j, \\ 0, i \neq j. \end{cases} \quad (2.5)$$

If the condition of orthonormality is fulfilled, then the components of the vector Y are determined as follows:

$$y_i = \Phi_i^T X, i = 1, \dots, n. \quad (2.6)$$

If *m* (*m*<*n*) components of the vector *Y* are determined, then the following formula can be used to estimate the vector *X* (at the same time, the unknown components of *Y* are replaced by preselected constants):

$$\hat{X}(m) = \sum_{i=1}^{m} y_i \Phi_i + \sum_{i=m+1}^{n} b_i \Phi_i. \quad (2.7)$$

Without restriction of commonality, we can assume that only the first m components of the vector y are calculated. If not all signs are used, then the vector *X* is represented with an error defined by the expression:

$$\Delta X(m) = X - \hat{X}(m) = X - \sum_{i=1}^{m} y_i \Phi_i - \sum_{i=m+1}^{n} b_i \Phi_i = \sum_{i=m+1}^{n} (y_i - b_i)\Phi_i, \quad (2.8)$$

where $\hat{X}$, ΔX are random vectors.

We will use the average value of the square ΔX as a criterion for measuring the efficiency of a subset consisting of m features:

$$\bar{\varepsilon}^2(m) = E\{\|\Delta X(m)\|^2\} = \sum_{i=m+1}^{n} E\{(y_i - b_i)^2\}. \quad (2.9)$$

Each set of base vectors and values of constants corresponds to some value $\varepsilon^2(m)$. It is necessary to choose them in such a way as to minimize $\overline{\varepsilon}^2(m)$. The optimal selection of b_i constants is performed as follows:

$$b_i = E\{y_i\} = \Phi_i^T E\{X\}. \tag{2.10}$$

It follows from this that it is necessary to replace y_i, which are not measured, by their mathematical expectations, which is easily proven by the example of minimizing $\overline{\varepsilon}^2(m)$ by b_i:

$$\frac{\partial}{\partial b_i} E\{(y_i - b_i)^2\} = -2[E\{y_i\}b_i] = 0. \tag{2.11}$$

Now the root mean square error can be written as follows:

$$\overline{\varepsilon}^2(m) = \sum_{i=m+1}^{n} E[(y_i - E\{y_i\})^2] = \sum_{i=m+1}^{n} \Phi_i^T \sum_x \Phi_i, \tag{2.12}$$

where Σ_x is the covariance matrix of the random vector X. The optimal choice of the matrix Φ will satisfy the condition

$$\sum_x \Phi_i = \lambda_i \Phi_i. \tag{2.13}$$

That is, the optimal basis vectors are the eigenvectors of the covariance matrix Σ_x. Thus, the minimum root mean square error is equal to

$$\overline{\varepsilon}^2(m)_{opt} = \sum_{i=m+1}^{n} \lambda_i \tag{2.14}$$

In pattern recognition tasks, the coefficients $y_1, y_2, \ldots, y_n$ of this distribution are considered as features representing the observed vector X.

These features have the following useful properties [9].

The effectiveness of each feature, that is, its utility from the point of view of the representation of X, is determined by the corresponding eigenvalue. If some feature, for example, y_i, is excluded from the schedule, then the root mean square error increases by λ_i. That is, if it is necessary to reduce the number of features, then the feature with the smallest eigenvalue must be eliminated first, etc. If the eigenvalues are renumbered in descending order $\lambda_1 > \lambda_2 > \ldots > \lambda_n > 0$, then the features should be ordered by importance in the same way.

The characteristics under consideration are mutually uncorrelated, that is, the covariance matrix of random of the vector Y is diagonal:

$$\sum_y = \Phi^T \sum_x \Phi = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix} = \Lambda. \tag{2.15}$$

In the case, when the random vector X is normally distributed, the features y_i are mutually independent.

The eigenvectors of the covariance matrix Σ_x give the smallest value of the root mean square error $\overline{\varepsilon}^2(m)$ among all orthonormal basis vectors. The ratio of the eigenvalue to the sum of all eigenvalues shows what proportion of the root mean square error is contributed by the exclusion of the corresponding eigenvector. Hence, the expression

$$\mu_i = \lambda_i \Big/ \sum_{j=1}^{n} \lambda_j = \lambda_i / tr \sum_x \tag{2.16}$$

can be used as a criterion for including or excluding the i-th eigenvector in the schedule.

The normalized vector Z is given by the expression

$$\sum_{i=1}^{n} \mu_i = 1; \ \|Z\| = 1. \tag{2.17}$$

Let Σ_z and i λ'_i be the covariance matrix of the normalized vector Z and its eigenvalues. Then

$$\sum_{i=1}^{n} \lambda'_i = tr \sum z = E\{Z^T Z\} = 1. \tag{2.18}$$

In other words, Σ_z i λ'_i are normalized eigenvalues. However, the transformation (2.17) must be justified from a physical point of view, since the statistical properties of the vector Z, including the covariance matrix, are completely different from the statistical properties of the vector X.

Thus, applying the Karhunen-Loeve decomposition to the vector Z gives completely different eigenvectors, and therefore completely different signs, than

applying the same decomposition to the original data.

III.Experimental researches

The study of nonlinearity was performed for the integrated microcircuit (IMs) of the TDA2593 synchroprocessor in the amount of 160 units on representative samples [8].

Complex dependences of quadratic nonlinearity on modulus and phase obtained by the difference frequency method separately for fit and defective chips. These dependences were transformed into components of *cosine Fc* [a_2 *(Uo)*] and *sine Fs* [a_2 *(Uo)*] and calculated by DKLE.

The number of necessary base vectors m allows to determine the dimension of the feature space with a given decomposition error $\varepsilon^2(m)$. The base vectors themselves are statistical standards of defective and suitable chips, and the decomposition coefficients are space coordinates.

For greater clarity, the Hilbert space is replaced by the Euclidean space. In the two-dimensional Euclidean orthonormal space (m=2, $\varepsilon^2(m)$=8%, n=100), intersections of the images of many suitable and defective chips, selected a priori based on expert data, were observed. At the same time, the DKLE coefficients of suitable microcircuits were mainly located near a circle of unit radius, and defective ones shifted to the middle of the circle.

When adding the third basis vector (m=3, $\varepsilon^2(m)$=6%, n=100), the orthonormal space of cosine vectors is transformed into a hemispher on the left, and sine vectors on the right. The points of the reflected suitable chips lie in the near-surface layer of a sphere of unit radius, and the reflections of different defective ones almost do not intersect. The decomposition of volumetri hemispheres made it possible to create a simple algorith for recognizing good and defective microcircuits.

In fact, when the DKLE error is reduced to 1%, the space of nonlinear basis vectors is an eight-dimensional ellipsoid, the reflection of which causes the appearance of folds and cusps in Euclidean space. This packing of space [10] does not interfere with the natural classification of defective and potentially unreliable microcircuits, at least, into 5 classes as balls inside the two main hemispheres with each hemisphere containing two or three small balls.

For the practical implementation of diagnostic information processing with complex biharmonic influence, it is proposed to use modern neural network technologies (multilayer perceptron, Kohonen maps, radial basis networks) [8,14,15]. The following algorithms were chosen for MLP training in the MATLAB environment [11,12]: Bayesian regularization or learning a function based on backpropagation of the error using Bayesian regularization; gradient descent backpropagation or gradient descent method; gradient descent method with adaptive backpropagation; backpropagation or Powell-Beale gradient method combined with Powell-Beale iterations; the elastic backpropagation method or the inverse elastic distribution method. International types of errors - MSE, MAE and others - were used as a criterion for assessing the accuracy of training.

For microcircuits, the classification is performed for the left and right hemispheres separately, so it is enough to perform the classification into two or three classes.

To verify the performance of partitioning in the volume after performing the classification, a set of test data consisting of three classes was generated: suitable chips (conventionally represented by points on the unit sphere), and two classes of defective chips that overlap in the space [13].

The result of perceptron classification in three-dimentional spase with different numbers of neurons in the inner layer is shown on Fig. 2-4.

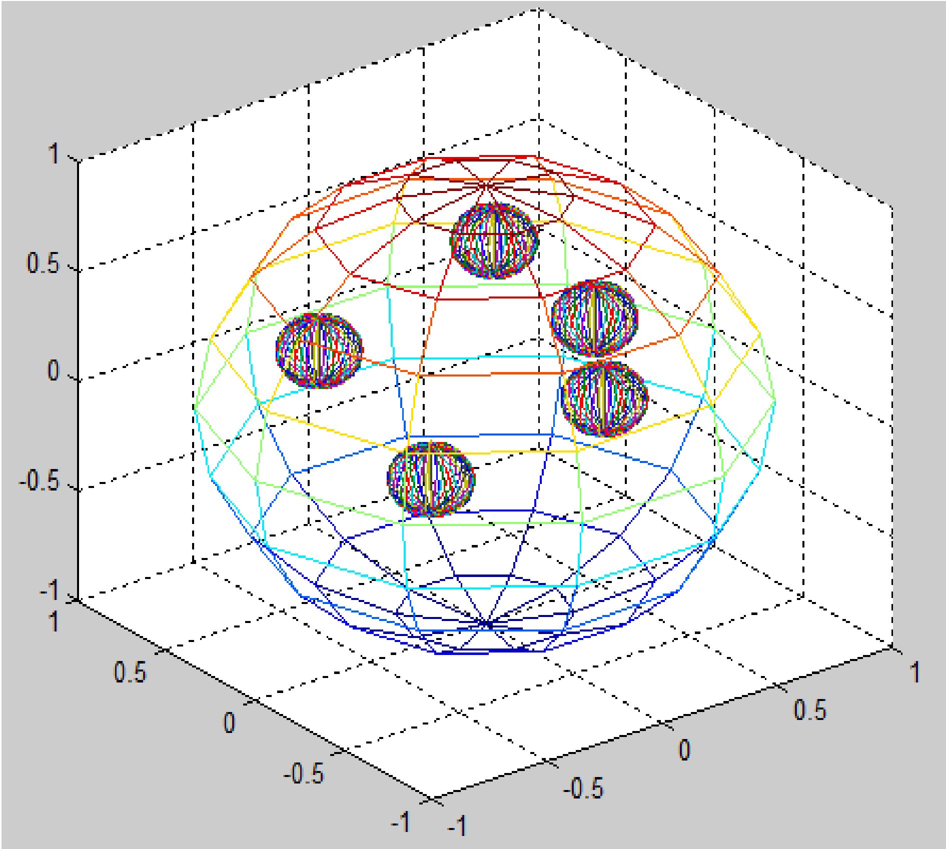


Fig. 1. Conditional representation of microcircuits classes in three-dimensional space

Confusion Matrix

Output classes	1	2	3	
1	39 26,0%	0 0%	0 0%	100% 0%
2	4 2,7%	46 30,7%	6 4,0%	82,1% 17,9%
3	7 4,7%	4 2,7%	44 29,3%	80,0% 20,0%
	78,0% 22,0%	92,0% 8,0%	88,0% 12,0%	86,0% 14,0%

Target Classes

Fig. 2. Classification result of the perceptron with 10 neurons in the inner layer

Confusion Matrix

Output Classes	1	2	3	
1	48 32,0%	0 0%	0 0%	100% 0%
2	0 0,0%	46 30,7%	4 2,7%	92,0% 8,0%
3	2 1,3%	4 2,7%	46 30,7%	68,5% 11,5%
	96,0% 4,0%	92,0% 8,0%	92,0% 8,0%	93,3% 6,7%

Target Classes

Fig. 3. Optimal classification result of the perceptron with 40 neurons in the inner layer

Confusion Matrix

Output Classes \ Target Classes	1	2	3	
1	47 31,3%	0 0%	0 0%	100% 0%
2	3 2,0%	45 30,0%	4 2,7%	86,5% 13,5%
3	0 0,0%	5 3,3%	46 30,7%	90,2% 11,5%
	94,0% 6,0%	90,0% 10,0%	92,0% 8,0%	92,0% 8,0%

Fig. 4. Classification result of the perceptron with 80 neurons in the inner layer

Therefore, to evaluate the success of the classification, it is necessary to determine whether the obtained classes overlap in volume. In this case, the probability that the resulting classes are independent is quite high.

The classification algorithm in three-dimensional space is proposed [13]. The advantage of the method is that it can be used when working with classes defined by more than three characteristics or vectors. For this, it must be scaled into an n-dimensional hyperspace. It is also possible to split objects on a plane if they are affected by two characteristics.

The perceptron training procedure using these algorithms is given in international types of errors were used to assess the accuracy of training results (Tables 1 and 2).

Table 1. Result of classification by projections of space (trainbr)

Number of classes	Classification accuracy, %	Training time, s
Two	99.23	24
Three	99.05	33

The perspective of the development of the division method consists in the application of fuzzy logic tools (since the allocation of volumetric areas, which are represented by classes, is quite conditional by its nature).

The increase in classification accuracy is associated with an increase in the number of neurons in the inner layer and leads to an improvement in the quality of classification, but only up to a certain limit, further complication of the network does not bring a positive effect.

Table 2. The best Results for Accuracy of the IMs Classification Using Multiplayer Perceptron

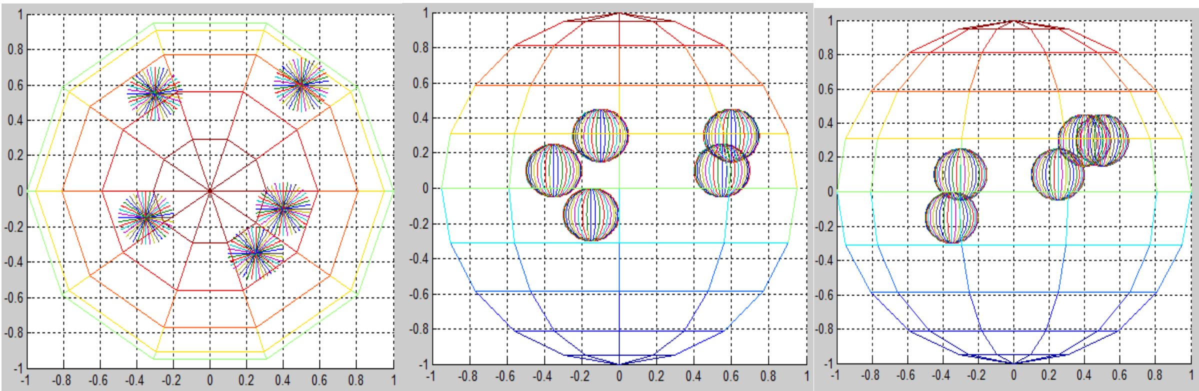
Number of classes	Activation Functions of Hidden Layers logsig- logsig- logsig					
	Error type, 10E-5				kep, unit	ttr, s
	MSE	SSE	MSEREG	MAE		
2	1.4	4200	1.8	1100	52	24
3	3	8000	3	2100	62	33

Thus the best results again were shown by neural network with activation function of each hidden layer. The data sample size is 164 IMs (Table 2).

Then for solving probability' problems a special type of neural network PNN (Probabilistic Neural Networks) is using [16-18]. PNN network architecture is based on the architecture of radial basis function network, but as a second layer uses so-called rival layer which calculates the probability of an input vector belonging to a particular class and compares the vector of the class whose probability of belonging to above [16].

Learning a probabilistic neural network is somehow easier than the method of reverse error propagation. The most important advantages of PNN networks are that the original value has a probabilistic meaning (and therefore is easier to interpret), and that the network learns quickly. When training such a network, time is spent almost only on

submitting training observations to it at the entrance, and the network works as fast as it can at all. The disadvantage of the network is its size, since it actually contains all the training data, requires a lot of memory and can work slowly.



a) b) c)

Fig. 5. The results of RBF network training: a) projection of the received DKLE matrices for microcircuits onto the XY plane; b) projection of the obtained DKLE matrices for microcircuits onto the YZ plane; c) projection of the obtained DKLE matrices for microcircuits onto the XZ plane

Table 3. Comparing the Best Results for All Training Types of Neural Network for Microcircuits

Neural Network Type	Classification Accuracy, %	Training Time, s	Classification Accuracy , %	Training Time, s
	Two basis vectors		Three basis vectors	
RBF pnn)	99.96	2	99,96	2
RBF (newrbe)	82.81	2	70,83	2
MLP	99.23	24	99.05	33

Classification and presorting of elements according to their physical and technical states were performed on neural networks in the MATLAB software package with its library Neural Network Toolbox [19]. For IMs presorting the probabilistic neural networks were used. In fact, the network is trained to assess the probability density function. According to the Bayesian statistics [20] to minimize error the model with the greatest probability density is being chosen. The resulting data for chips confirm the highest accuracy 99.96% and speed training time 2 hours at using the probabilistic (pnn) RBF networks compared with the neural networks with the radial- basis elements and given zero error (newrbe) (Table 3).

Further study was carried out on the self-organized maps [8,11]. In this way, the neurons of the active sphere have a regular structure. Such cards are often used for clustering graphic images and sound signals, as well as for processing rich information. A self-organized (Kohonen Map) network (excluding teachers) implements a neural network algorithm known as analogue *K* medium. At each step of the training input vector network served and sought another neuron whose weight differs least from this vector.

Found decla declared the winner neuron and its weights vector *W* is updated as follows:

$$w_i(k+1)= w_i(k)+ \eta(x_n-w_i(k))+aw_i(k)(1-||w_i||^2), \quad (3.1)$$

which parameter η is responsible for setting the rate of learning and change its value in the interval (0,1). All training vectors are

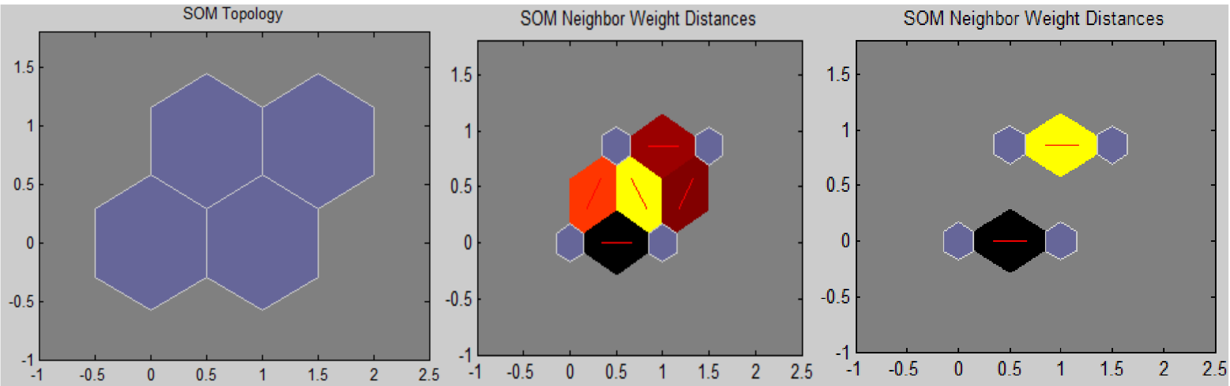
processed one by one until they will fail to stabilize or other condition stops.

Table 4. Dependence of the Accuracy Classification on the Self-Orgenizerd Map Setting Parameters

Topology	Hextop			
Distance type between adjacent clasters	Linkdist	Dist	Mandist	Boxdist
Correctly classified samples, %	83.33	83.5	61	70.67
Topology	Gridtop			
Distance type between adjacent clasters	Linkdist	Dist	Mandist	Boxdist
Correctly classified samples, %	61	61	83.67	83.33
Number of training periods	200			
Average training time,s	2			

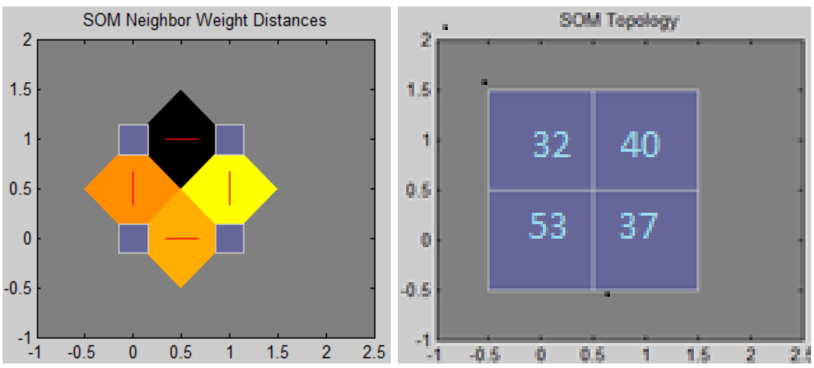
The results in the Table 4 have the best performance with the “gridtop” topology at using the distance between the “mandist” clusters, and moreover impact on the classification accuracy step parameter used to estimate distance between the neighboring clusters is not significant (Fig. 6 and 7).

Thus, the best accuracy (83,67%) can be achieved for gridtop topology with mandist distances and steps=30 step parameter (Table 4 and 5).



a) b) c)

Fig. 6. The principle of forming distances between adjacent clusters: a) hextop topology; b) the principle of forming distances for functions dist, linkdist and boxdist; c) for mandist



a) b)

Fig. 7. The principle of forming distances between adjacent clusters: a) for gridtop topology with mandist distances; b) classification results on the Kohonen map

Table 5. Influence of the Step Parameters on the Classification Accuracy

“Gridtop” Topology, Distance between “Mandist” Clusters							
Step size	10	30	50	70	80	90	100
Probability of Classification	0.61	0.84	0.83	0.59	0.6	0.61	0.837

The applied technique traingd (gradient descent backpropagation) and/or penalty functions gave good results, but worse than RBF.

Therefore, the best result of chip rejection is provided by RBF neural networks (pnn) when trained with a teacher.

Conclusion

The use of discrete Karhunen-Loeve expansion for compressing a priori information about the microcircuits technical status was theoretically justified and practically implemented, which made it possible to carry out elements individual sorting and statistical evaluation of the IT structure components batch. Determination of components technical status using MLP, self-organized (Kohonen Map) and RBF neural networks in the MATLAB environment was tested. The best result was obtained for probabilistic RBF-neural networks. The neural network technology usage is proposed to improve the electrophysical methods of diagnosing the IT structure components, which significantly increased the accuracy and speed of diagnosis.

References

1. *Strategy for Artificial Intelligence Development in Ukraine: monograph* / A. Shevchenko, S. Baranovsky, O. Bilokobylskyi and others. [Under the general editorship of A. Shevchenko] (2023). Kyiv: IAIP.
2. Rolik, O.I., Telenyk, C.F., Yasochka, M.V. (2018). *Corporate IT – infrastructure management*. Kyiv: Naukova Dumka.
3. *Methods and latest approaches to design, management and application of high-performance IT infrastructures: monograph* / Yu.V. Boyko, V. M., Volokhov, M. M. Glibovets, S. V. Yershov, S. L. Kryviy, S. D. Pohorily, O. I. Rolik, S. F. Telenyk, M. V. Yasochka. [Edited by prof. A. V. Anisimov] (2016). Higher Vocational Training Center "Kyiv University".
4. Volodarskyi, E.T., Kuharchuk, V.V., Podzharenko, V.O., Serdjuk, G.B. (2001). *Metrological Support for Measuring and Checking / Textbook.* Vinnytsia: Veles.
5. Usatenko, V.G. Serdjuk, G.B., Sazonov, S.M. (1993). *Unrealible and Unstable Microcircuits’ regection method.* Patent N1056 for an invention (Ukraine). - G 01 R 31/27.
6. Castrillon-Candas, J., Amaratunga, K. (2002). *Fast estimation of continuous Karhunen-Loeve eingenfunctions using wavelets.* – IEEE: IEEE Signal Processing Society.
7. Julio E. Castrillon-Candas, Jun Li, Victor Eijkhout. (2012). *A discrete adapted hierarchical basis solver for radial basis function interpolation* / BIT Numerical Mathematics, Springer.
8. Telenyk S., Savchuk O., Pocrovskyi E., Morgal O. Krivenko, Latash I. (2017). *On Searching of Type and Optimizing of Neural Network in Electric Component Diagnosis* // ICACIT’2017. Cracow, Poland.
9. Fucunaga K. (1990). *Introduction to Pattern Recognition Statistical Theory-Second Edition.* Computer Science. URL: https:// www.semanticshoolar.org/paper/ introduction-to-statistic-Edition-Fucunaga/ 97de254c563134d233fd7f1ee40f68a8adc035c
10. Henry Kohn, Abhinav Kumar, Stephen Miller, Danylo Radchenko, Maryna Vyazovska (2022). *Universal Optimality in E8 and Leach Lattices and Interpolation Formulas.* / Ann. of Math. (2) 196 (3): 983 – 1082. https://doi.org/10.4007/annals.2022.196.3.3
11. Dyakonov V., Kruglov V. (2001). *MATLAB Math Expantion Packs.* Special Handbook.
12. Yakimenko O. (2011). *Engineering computations and modeling in MATLAB/Simulink*. – Monterey, California: American Institute of Aeronautics and Astronautics, INC.
13. Savchuk O. V., Zhukovetskyi B. V. (2015). *To the Question of Electro-radio Products Classification and Rejection* / InfoCom'15. Kyiv, Ukraine.
14. Haikin, S. (2006) *Neural Networks: Full Course.2nd edition*. – M.: Williams.
15. Borovikov V.P. (2008). Neural Networks STATISTICA Neural-Networks: Metodology and Technology of modern data analysis. Hotline. - Telecom.
16. Broomhead, D. S.; Lowe, David (1988). *Radial basis functions, multi- variable functional interpolation and adaptive networks (Technical report)*. *RSRE*. 4148. Archived from the original on April 9, 2013.

17. Ismayilova, A., Ismayilov, M. (2023). On the universal approximation property of radial basis function neural networks. *Ann Math Artif Intell*. URL: https://doi.org/10.1007/s10472-023-09901-x

18. Shymkovych, V., Telenyk, S. & Kravets, P. (2021). Hardware implementation of radial-basis neural networks with Gaussian activation functions on FPGA. *Neural Computing and Applications,* 33, 9467–9479. URL: https://doi.org/10.1007/s00521-021-05706-3

19. Howard Demuth Mark Beal. *Neural Network Toolbox For Use with MATLAB*. URL: http://cda.psych.uiuc.edu/matlab_pdf/nnet.pdf

20. Pysarenko, A., Tyshenko, D. (2017). *The Model of Vehicle Diagnostic Subsystem Based of Bayesian Network* / Infocom'2017. Kyiv: Engineering.

The article has been sent to the editors 30.11.23
After processing 25.12.23
Submitted for printing 20.03.24